ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGATGGGAAAAAAT

TCGGTTAAGGCCAGGGGGAAAGAAGAAGTACAAGCTAAAGCACATCGTATGGGCAA

GCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGC

TGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAGGAGCT

TCGATCACTATACAACACAGTAGCAACCCTCTATTGTGTGCACCAGCGGATCGAGA

TCAAGGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAGTCCAAG

AAGAAGGCCCAGCAGGCAGCAGCTGACACAGGACACAGCAATCAGGTCAGCCAAAA

TTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTA

GAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTG

ATACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAGGACCTGAACACGAT

GTTGAACACCGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCA

ATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCA

CCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACTACTAGTACCCT

TCAGGAACAAATAGGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAGATCT

ACAAGAGGTGGATAATCCTGGGATTGAACAAGATCGTGAGGATGTATAGCCCTACC

AGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCG

GTTCTATAAAACTCTAAGAGCTGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGA

CAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACCATCCTGAAGGCT

CTCGGCCCAGCGGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGG

ACCCGGCCATAAGGCAAGAGTTTTGGCCGAGGCGATGAGCCAGGTGACGAACTCGG

CGACCATAATGATGCAGAGAGGCAACTTCCGGAACCAGCGGAAGATCGTCAAGTGC

TTCAATTGTGGCAAAGAAGGGCACACCGCCAGGAACTGCCGGGCCCCCCGGAAGAA

GGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTACTGAGAGAC

FIG. IA

AGGCTAATTTTTTAGGGAAGATCTGGCCTTCCTACAAGGGAAGGCCAGGGAATTTT

CTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCTGGGGT

AGAGACAACAACTCCCCCTCAGAAGCAGGAGCCGATAGACAAGGAACTGTATCCTT

TAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAGTAAGGATCGGGG

GGCAACTCAAGGAAGCGCTGCTCGATACAGGAGCAGATGATACAGTATTAGAAGAA

ATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGGATCGGGGGCTTCAT

CAAGGTGAGGCAGTACGACCAGATACTCATAGAAATCTGTGGACATAAAGCTATAG

GTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACC

CAGATCGGCTGCACCTTGAACTTCCCCATCAGCCCTATTGAGACGGTGCCCGTGAA

GTTGAAGCCGGGGATGGACGGCCCCAAGGTCAAGCAATGGCCATTGACGAAAGAGA

AGATCAAGGCCTTAGTCGAAATCTGTACAGAGATGGAGAAGGAAGGGAAGATCAGC

AAGATCGGGCCTGAGAACCCCTACAACACTCCAGTCTTCGCAATCAAGAAGAAGGA

CAGTACCAAGTGGAGAAAGCTGGTGGACTTCAGAGAGCTGAACAAGAGAACTCAGG

ACTTCTGGGAAGTTCAGCTGGGCATCCCACATCCCGCTGGGTTGAAGAAGAAGAAG

TCAGTGACAGTGCTGGATGTGGGTGATGCCTACTTCTCCGTTCCCTTGGACGAGGA

CTTCAGGAAGTACACTGCCTTCACGATACCTAGCATCAACAACGAGACACCAGGCA

TCCGCTACCAGTACAACGTGCTGCCACAGGGATGGAAGGGATCACCAGCCATCTTT

CAAAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAAAACCCAGACATCGT

GATCTATCAGTACATGGACGACCTCTACGTAGGAAGTGACCTGGAGATCGGGCAGC

ACAGGACCAAGATCGAGGAGCTGAGACAGCATCTGTTGAGGTGGGGACTGACCACA

CCAGACAAGAAGCACCAGAAGGAACCTCCCTTCCTGTGGATGGGCTACGAACTGCA

TCCTGACAAGTGGACAGTGCAGCCCATCGTGCTGCCTGAGAAGGACAGCTGGACTG

TGAACGACATACAGAAGCTCGTGGGCAAGTTGAACTGGGCAAGCCAGATCTACCCA

GGCATCAAAGTTAGGCAGCTGTGCAAGCTGCTTCGAGGAACCAAGGCACTGACAGA

FIG. IB

AGTGATCCCACTGACAGAGGAAGCAGAGCTAGAACTGGCAGAGAACCGAGAGATCC

TGAAGGAGCCAGTACATGGAGTGTACTACGACCCAAGCAAGGACCTGATCGCAGAG

ATCCAGAAGCAGGGGCAAGGCCAATGGACCTACCAAATCTACCAGGAGCCCTTCAA

GAACCTGAAGACAGGCAAGTACGCAAGGATGAGGGGTGCCCACACCAACGATGTGA

AGCAGCTGACAGAGGCAGTGCAGAAGATCACCACAGAGAGCATCGTGATCTGGGGC

AAGACTCCCAAGTTCAAGCTGCCCATACAGAAGGAGACATGGGAGACATGGTGGAC

CGAGTACTGGCAAGCCACCTGGATCCCTGAGTGGGAGTTCGTGAACACCCCTCCCT

TGGTGAAACTGTGGTATCAGCTGGAGAAGGAACCCATCGTGGGAGCAGAGACCTTC

TACGTGGATGGGGCAGCCAACAGGGAGACCAAGCTGGGCAAGGCAGGCTACGTGAC

CAACCGAGGACGACAGAAAGTGGTGACCCTGACTGACACCACCAACCAGAAGACTG

AGCTGCAAGCCATCTACCTAGCTCTGCAAGACAGCGGACTGGAAGTGAACATCGTG

ACAGACTCACAGTACGCACTGGGCATCATCCAAGCACAACCAGACCAATCCGAGTC

AGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGTGTACCTGG

CATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTC

AGTGCTGGGATCCGGAAGGTGCTGTTCCTGGACGGGATCGATAAGGCCCAAGATGA

ACATGAGAAGTACCACTCCAACTGGCGCGCTATGGCCAGCGACTTCAACCTGCCAC

CTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAA

GCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAGCTGGACTGCACGCA

CCTGGAGGGGAAGGTGATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAG

CAGAAGTTATCCCTGCTGAAACTGGGCAGGAAACAGCATATTTTCTTTTAAAATTA

GCAGGAAGATGGCCAGTAAAAACAATACACACGGACAACGGAAGCAACTTCACTGG

TGCTACGGTTAAGGCCGCCTGTTGGTGGGCGGGAATCAAGCAGGAATTTGGAATTC

CCTACAATCCCCAATCGCAAGGAGTCGTGGAGAGCATGAACAAGGAGCTGAAGAAG

ATCATCGGACAAGTGAGGGATCAGGCTGAGCACCTGAAGACAGCAGTGCAGATGGC

FIG. IC

AGTGTTCATCCACAACTTCAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGG

AAAGGATCGTGGACATCATCGCCACCGACATCCAAACCAAGGAGCTGCAGAAGCAG

ATCACCAAGATCCAGAACTTCCGGGTGTACTACCGCGACAGCCGCAACCCACTGTG

GAAGGGACCAGCAAAGCTCCTCTGGAAGGGAGAGGGGGCAGTGGTGATCCAGGACA

ACAGTGACATCAAAGTGGTGCCAAGGCGCAAGGCCAAGATCATCCGCGACTATGGA

AAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGAACCT

GGAAGAGCCTGGTGAAGCACCATATG (SEQUENCE ID NO:1)

FIG. ID

```
>wildtype       TGTACAGAGA TGGAAAAGGA AGGGAAAATT TCAAAAATTG
>mutated        TGTACAGAGA TGGAGAAGGA AGGGAAGATC AGCAAGATCG
 #1             ..........................................
                          *           * * ***  * *

>wildtype       GGCCTGAAAA TCCATACAAT ACTCCAGTAT TTGCCATAAA
>mutated        GGCCTGAGAA CCCCTACAAC ACTCCAGTCT TCGCAATCAA
 #41            ..........................................
                    * * *   *       *   * * *

>wildtype       GAAAAAAGAC AGTACTAAAT GGAGAAAATT AGTAGATTTC
>mutated        GAAGAAGGAC AGTACCAAGT GGAGAAAGCT GGTGGACTTC
 #81            ..........................................
                  * *     * *       **  * * *

>wildtype       AGAGAACTTA ATAAGAGAAC TCAAGACTTC TGGGAAGTTC
>mutated        AGAGAGCTGA ACAAGAGAAC TCAGGACTTC TGGGAAGTTC
 #121           ..........................................
                    * * *         *

>wildtype       AATTAGGAAT ACCACATCCC GCAGGGTTAA AAAAGAAAAA
>mutated        AGCTGGGCAT CCCACATCCC GCTGGGTTGA AGAAGAAGAA
 #161           ..........................................
                 ** * *  *         *   *  *   *

>wildtype       ATCAGTAACA GTACTGGATG TGGGTGATGC ATATTTTTCA
>mutated        GTCAGTGACA GTGCTGGATG TGGGTGATGC CTACTTCTCC
 #201           ..........................................
                *    *     *                 * * * *

>wildtype       GTTCCCTTAG ATGAAGACTT CAGGAAATAT ACTGCATTTA
>mutated        GTTCCCTTGG ACGAGGACTT CAGGAAGTAC ACTGCCTTCA
 #241           ..........................................
                       * * *           * *    * *

>wildtype       CCATACCTAG TATAAACAAT GAGACACCAG GGATTAGATA
>mutated        CGATACCTAG CATCAACAAC GAGACACCAG GCATCCGCTA
 #281           ..........................................
                *       * *  *           *  ** *

>wildtype       TCAGTACAAT GTGCTTCCAC AGGGATGGAA AGGATCACCA
>mutated        CCAGTACAAC GTGCTGCCAC AGGGATGGAA GGGATCACCA
 #321           ..........................................
                *     *    *                *

>wildtype       GCAATATTCC AAAGTAGCAT GACAAAAATC TTAGAGCCTT
>mutated        GCCATCTTTC AAAGCAGCAT GAGCAAGATC CTGGAGCCCT
 #361           ..........................................
                  * * *     *       * *  * *    *

>wildtype       TTAGAAAACA AAATCCAGAC ATAGTTATCT ATCAATACAT
>mutated        TCCGCAAGCA AAACCCAGAC ATCGTGATCT ATCAGTACAT
 #401           ..........................................
                ** * *     *         * *        *
```

FIG. 2A

```
>wildtype       GGATGATTTG TATGTAGGAT CTGACTTAGA AATAGGGCAG
>mutated        GGACGACCTC TACGTAGGAA GTGACCTGGA GATCGGGCAG
 #441           ..........................................
                   *  **  *   *      *  *    *  *   *  *

>wildtype       CATAGAACAA AAATAGAGGA GCTGAGACAA CATCTGTTGA
>mutated        CACAGGACCA AGATCGAGGA GCTGAGACAG CATCTGTTGA
 #481           ..........................................
                  *  *  *   *  *              *

>wildtype       GGTGGGGACT TACCACACCA GACAAAAAAC ATCAGAAAGA
>mutated        GGTGGGGACT GACCACACCA GACAAGAAGC ACCAGAAGGA
 #521           ..........................................
                        *                 *  *   *    *

>wildtype       ACCTCCATTC CTTTGGATGG GTTATGAACT CCATCCTGAT
>mutated        ACCTCCCTTC CTGTGGATGG GCTACGAACT GCATCCTGAC
 #561           ..........................................
                     *     *        *  *     *        *

>wildtype       AAATGGACAG TACAGCCTAT AGTGCTGCCA GAAAAAGACA
>mutated        AAGTGGACAG TGCAGCCCAT CGTGCTGCCT GAGAAGGACA
 #601           ..........................................
                  *        *   *   *        *  *  *

>wildtype       GCTGGACTGT CAATGACATA CAGAAGTTAG TGGGGAAATT
>mutated        GCTGGACTGT GAACGACATA CAGAAGCTCG TGGGCAAGTT
 #641           ..........................................
                        *  *              *  *     *  *

>wildtype       GAATTGGGCA AGTCAGATTT ACCCAGGGAT TAAAGTAAGG
>mutated        GAACTGGGCA AGCCAGATCT ACCCAGGCAT CAAAGTTAGG
 #681           ..........................................
                   *        *   *        *   *    *

>wildtype       CAATTATGTA AACTCCTTAG AGGAACCAAA GCACTAACAG
>mutated        CAGCTGTGCA AGCTGCTTCG AGGAACCAAG GCACTGACAG
 #721           ..........................................
                  **  *  *   *  *  *           *     *

>wildtype       AAGTAATACC ACTAACAGAA GAAGCAGAGC TAGAACTGGC
>mutated        AAGTGATCCC ACTGACAGAG GAAGCAGAGC TAGAACTGGC
 #761           ..........................................
                    *  *     *    *

>wildtype       AGAAAACAGA GAGATTCTAA AAGAACCAGT ACATGGAGTG
>mutated        AGAGAACCGA GAGATCCTGA AGGAGCCAGT ACATGGAGTG
 #801           ..........................................
                   *  *        *  *  *  *

>wildtype       TATTATGACC CATCAAAAGA CTTAATAGCA GAAATACAGA
>mutated        TACTACGACC CAAGCAAGGA CCTGATCGCA GAGATCCAGA
 #841           ..........................................
                  *  *      ***  *   *  *  *     *  *
```

FIG. 2B

```
>wildtype    AGCAGGGGCA AGGCCAATGG ACATATCAAA TTTATCAAGA
>mutated     AGCAGGGGCA AGGCCAATGG ACCTACCAAA TCTACCAGGA
 #881        ..........................................
                                    *  *     *  *  *

>wildtype    GGCATTTAAA AATCTGAAAA CAGGAAAATA TGCAAGAATG
>mutated     GCCCTTCAAG AACCTGAAGA CAGGCAAGTA CGCAAGGATG
 #921        ..........................................
               *  *  *   *     *     *  *   *     *

>wildtype    AGGGGTGCCC ACACTAATGA TGTAAAACAA TTAACAGAGG
>mutated     AGGGGTGCCC ACACCAACGA TGTGAAGCAG CTGACAGAGG
 #961        ..........................................
                        *  *     *  *  *  *  *

>wildtype    CAGTGCAAAA AATAACCACA GAAAGCATAG TAATATGGGG
>mutated     CAGTGCAGAA GATCACCACA GAGAGCATCG TGATCTGGGG
 #1001       ..........................................
                  *  *  *     *     *  *  *

>wildtype    AAAGACTCCT AAATTTAAAC TGCCCATACA AAAGGAAACA
>mutated     CAAGACTCCC AAGTTCAAGC TGCCCATACA GAAGGAGACA
 #1041       ..........................................
             *     *  *  *  *              *     *

>wildtype    TGGGAAACAT GGTGGACAGA GTATTGGCAA GCCACCTGGA
>mutated     TGGGAGACAT GGTGGACCGA GTACTGGCAA GCCACCTGGA
 #1081       ..........................................
                  *           *     *

>wildtype    TTCCTGAGTG GGAGTTTGTT AATACCCCTC CTTTAGTGAA
>mutated     TCCCTGAGTG GGAGTTCGTG AACACCCCTC CCTTGGTGAA
 #1121       ..........................................
              *              *  *  *           *  *

>wildtype    ATTATGGTAC CAGTTAGAGA AAGAACCCAT AGTAGGAGCA
>mutated     ACTGTGGTAT CAGCTGGAGA AGGAACCCAT CGTGGGAGCA
 #1161       ..........................................
              *  *     *     *  *     *           *  *

>wildtype    GAAACCTTCT ATGTAGATGG GGCAGCTAAC AGGGAGACTA
>mutated     GAGACCTTCT ACGTGGATGG GGCAGCCAAC AGGGAGACCA
 #1201       ..........................................
               *           *  *              *        *

>wildtype    AATTAGGAAA AGCAGGATAT GTTACTAATA GAGGAAGACA
>mutated     AGCTGGGCAA GGCAGGCTAC GTGACCAACC GAGGACGACA
 #1241       ..........................................
              **  *   *     *     *  *   *  *  **     *

>wildtype    AAAAGTTGTC ACCCTAACTG ACACAACAAA TCAGAAGACT
>mutated     GAAAGTGGTG ACCCTGACTG ACACCACCAA CCAGAAGACT
 #1281       ..........................................
             *    *  *     *           *  *  *
```

FIG. 2C

```
>wildtype       GAGTTACAAG CAATTTATCT AGCTTTGCAG GATTCGGGAT
>mutated        GAGCTGCAAG CCATCTACCT AGCTCTGCAA GACAGCGGAC
 #1321          ..........................................
                   * *      * * *       *    *   ****   *

>wildtype       TAGAAGTAAA CATAGTAACA GACTCACAAT ATGCATTAGG
>mutated        TGGAAGTGAA CATCGTGACA GACTCACAGT ACGCACTGGG
 #1361          ..........................................
                 *     *      * *             *   *  * *

>wildtype       AATCATTCAA GCACAACCAG ATCAAAGTGA ATCAGAGTTA
>mutated        CATCATCCAA GCACAACCAG ACCAATCCGA GTCAGAGCTG
 #1401          ..........................................
                *     *                 *   ***   *     * *

>wildtype       GTCAATCAAA TAATAGAGCA GTTAATAAAA AAGGAAAAGG
>mutated        GTGAACCAGA TCATCGAGCA GCTGATCAAG AAGGAGAAAG
 #1441          ..........................................
                  *  *  *   *  *       * *  *  *      *  *

>wildtype       TCTATCTGGC ATGGGTACCA GCACACAAAG GAATTGGAGG
>mutated        TGTACCTGGC ATGGGTACCA GCACACAAAG GAATTGGAGG
 #1481          ..........................................
                 *  *

>wildtype       AAATGAACAA GTAGATAAAT TAGTCAGTGC TGGAATCAGG
>mutated        AAATGAACAA GTAGATAAAT TAGTCAGTGC TGGGATCCGG
 #1521          ..........................................
                                                  *   *

>wildtype       AAAGTACTAT TTTTAGATGG AATAGATAAG GCCCAAGATG
>mutated        AAGGTGCTGT TCCTGGACGG GATCGATAAG GCCCAAGATG
 #1561          ..........................................
                  *  *  *   ** *  *   *  *

>wildtype       AACATGAGAA ATATCACAGT AATTGGAGAG CAATGGCTAG
>mutated        AACATGAGAA GTACCACTCC AACTGGCGCG CTATGGCCAG
 #1601          ..........................................
                          *  *   ***   *   * *   *     *

>wildtype       TGATTTTAAC CTGCCACCTG TAGTAGCAAA AGAAATAGTA
>mutated        CGACTTCAAC CTGCCACCTG TAGTAGCAAA AGAAATAGTA
 #1641          ..........................................
                *  *  *

>wildtype       GCCAGCTGTG ATAAATGTCA GCTAAAAGGA GAAGCCATGC
>mutated        GCCAGCTGTG ATAAATGTCA GCTAAAAGGA GAAGCCATGC
 #1681          ..........................................

>wildtype       ATGGACAAGT AGACTGTAGT CCAGGAATAT GGCAACTAGA
>mutated        ATGGACAAGT AGACTGTAGT CCAGGAATAT GGCAGCTGGA
 #1721          ..........................................
                                                    *  *
```

FIG. 2D

```
>wildtype    TTGTACACAT TTAGAAGGAA AAGTTATCCT GGTAGCAGTT
>mutated     CTGCACGCAC CTGGAGGGGA AGGTGATCCT GGTAGCAGTT
 #1761       ..........................................
             * * * * * * * * * *

>wildtype    CATGTAGCCA GTGGATATAT AGAAGCAGAA GTTATTCCAG
>mutated     CATGTAGCCA GTGGATATAT AGAAGCAGAA GTTATCCCTG
 #1801       ..........................................
                                                 * *

>wildtype    CAGAAACAGG GCAGGAAACA GCATATTTTC TTTTAAAATT
>mutated     CTGAAACTGG GCAGGAAACA GCATATTTTC TTTTAAAATT
 #1841       ..........................................
              *     *

>wildtype    AGCAGGAAGA TGGCCAGTAA AAACAATACA TACAGACAAT
>mutated     AGCAGGAAGA TGGCCAGTAA AAACAATACA CACGGACAAC
 #1881       ..........................................
                                           * *    *

>wildtype    GGCAGCAATT TCACCAGTGC TACGGTTAAG GCCGCCTGTT
>mutated     GGAAGCAACT TCACTGGTGC TACGGTTAAG GCCGCCTGTT
 #1921       ..........................................
               *     *     **

>wildtype    GGTGGGCGGG AATCAAGCAG GAATTTGGAA TTCCCTACAA
>mutated     GGTGGGCGGG AATCAAGCAG GAATTTGGAA TTCCCTACAA
 #1961       ..........................................

>wildtype    TCCCCAAAGT CAAGGAGTAG TAGAATCTAT GAATAAAGAA
>mutated     TCCCCAATCG CAAGGAGTCG TGGAGAGCAT GAACAAGGAG
 #2001       ..........................................
                  ***       *   *  ****     *  *  *

>wildtype    TTAAAGAAAA TTATAGGACA GGTAAGAGAT CAGGCTGAAC
>mutated     CTGAAGAAGA TCATCGGACA AGTGAGGGAT CAGGCTGAGC
 #2041       ..........................................
             * *     *   *   *     *   *   *          *

>wildtype    ATCTTAAGAC AGCAGTACAA ATGGCAGTAT TCATCCACAA
>mutated     ACCTGAAGAC AGCAGTGCAG ATGGCAGTGT TCATCCACAA
 #2081       ..........................................
              * *           *  *        *

>wildtype    TTTTAAAAGA AAAGGGGGGA TTGGGGGGTA CAGTGCAGGG
>mutated     CTTCAAAAGA AAAGGGGGGA TTGGGGGGTA CAGTGCAGGG
 #2121       ..........................................
             * *

>wildtype    GAAAGAATAG TAGACATAAT AGCAACAGAC ATACAAACTA
>mutated     GAAAGGATCG TGGACATCAT CGCCACCGAC ATCCAAACCA
 #2161       ..........................................
                * *   *     *   *   *   *     *     *
```

FIG. 2E

```
>wildtype       AAGAATTACA AAAACAAATT ACAAAAATTC AAAATTTTCG
>mutated        AGGAGCTGCA GAAGCAGATC ACCAAGATCC AGAACTTCCG
 #2201          ..........................................
                 *  **  *   *  *  *  *   *  *  *   *  *  *

>wildtype       GGTTTATTAC AGGGACAGCA GAAATCCACT TTGGAAAGGA
>mutated        GGTGTACTAC CGCGACAGCC GCAACCCACT GTGGAAGGGA
 #2241          ..........................................
                   *  *    * *      *  *  *      *     *

>wildtype       CCAGCAAAGC TCCTCTGGAA AGGTGAAGGG GCAGTAGTAA
>mutated        CCAGCAAAGC TCCTCTGGAA GGGAGAGGGG GCAGTGGTGA
 #2281          ..........................................
                                     *  *  *        *  *

>wildtype       TACAAGATAA TAGTGACATA AAAGTAGTGC CAAGAAGAAA
>mutated        TCCAGGACAA CAGTGACATC AAAGTGGTGC CAAGGCGCAA
 #2321          ..........................................
                 *  *  *  *        *       *        **  *

>wildtype       AGCAAAGATC ATTAGGGATT ATGGAAAACA GATGGCAGGT
>mutated        GGCCAAGATC ATCCGCGACT ATGGAAAACA GATGGCAGGT
 #2361          ..........................................
                *  *        **  *  *

>wildtype       GATGATTGTG TGGCAAGTAG ACAGGATGAG GATTAGAACA
>mutated        GATGATTGTG TGGCAAGTAG ACAGGATGAG GATTAGAACC
 #2401          ..........................................
                                                        *

>wildtype       TGGAAAAGTT TAGTAAAACA CCATATG
>mutated        TGGAAGAGCC TGGTGAAGCA CCATATG
 #2441          ...........................................
                    *  **  *  *  *
```

FIG. 2F

```
SIV gag   ----------------------------------------
 #1       ........................................
          ATGGGCGTGAGAAACTCCGTCTTGTCAGGGAAGAAAGCAG

SIV gag   ----------------------------------------
 #41      ........................................
          ATGAATTAGAAAAAATTAGGCTACGACCCAACGGAAAGAA

SIV gag   ----------------------------------------
 #81      ........................................
          AAAGTACATGTTGAAGCATGTAGTATGGGCAGCAAATGAA

SIV gag   ----------------------------------------
 #121     ........................................
          TTAGATAGATTTGGATTAGCAGAAAGCCTGTTGGAGAACA

SIV gag   ----------------------------------------
 #161     ........................................
          AAGAAGGATGTCAAAAAATACTTTCGGTCTTAGCTCCATT

SIV gag   ----------------------------------------
 #201     ........................................
          AGTGCCAACAGGCTCAGAAAATTTAAAAAGCCTTTATAAT

SIV gag   ----------------------------------------
 #241     ........................................
          ACTGTCTGCGTCATCTGGTGCATTCACGCAGAAGAGAAAG

SIV gag   ----------------------------------------
SIVgagDX..                           ----------
 #281     ........................................
          TGAAACACACTGAGGAAGCAAAACAGATAGTGCAGAGACA

SIV gag   --------------------A--A-----T-----A--A
SIVgagDX..--------------------C--C-----C-----G--G
 #321     ........................................
          CCTAGTGGTGGAAACAGGAACMACMGAAACYATGCCRAAR

SIV gag   --AAG-A-------------------------------
SIVgagDX..--CTC-C-------------------------------
 #361     ........................................
          ACMWSTMGACCAACAGCACCATCTAGCGGCAGAGGAGGAA
```

FIG. 4A

```
SIV gag    -T----------A--A--A----------T--------
SIVgagDX..-C----------G--G--C----------C--------
 #401       .....................................
           AYTACCCAGTACARCARATMGGTGGTAACTAYGTCCACCT

SIV gag    ----T-AAG---------AT-A--T--C-----A--AT--
SIVgagDX..----C-GTC---------CC-G--C--T-----C--GC--
 #441       .....................................
           GCCAYTRWSCCCGAGAACMYTRAAYGCYTGGGTMAARYTG

SIV gag    --A-----A-----A--T---------------------
SIVgagDX..--C-----G-----G--C---------------------
 #481       .....................................
           ATMGAGGARAAGAARTTYGGAGCAGAAGTAGTGCCAGGAT

SIV gag    -T-----------------------------T-----T--
SIVgagDX..-C-----------------------------C-----C--
 #521       .....................................
           TYCAGGCACTGTCAGAAGGTTGCACCCCCTAYGACATYAA

SIV gag    T------T-A--T--T--G----------A---------
SIVgagDX..C------C-G--C--C--T----------G---------
 #561       .....................................
           YCAGATGYTRAAYTGYGTKGGAGACCATCARGCGGCTATG

SIV gag    -----T---A-A--T--T--A-------------------
SIVgagDX..-----C---C-T--C--C--C-------------------
 #601       .....................................
           CAGATYATCMGWGAYATYATMAACGAGGAGGCTGCAGATT

SIV gag    -------------------------------------
SIVgagDX..-------------------------------------
 #641       .....................................
           GGGACTTGCAGCACCCACAACCAGCTCCACAACAAGGACA

SIV gag    ----------------------T--T--------A--T
SIVgagDX..----------------------C--C--------C--C
 #681       .....................................
           ACTTAGGGAGCCGTCAGGATCAGAYATYGCAGGAACMACY

SIV gag    AGT-----A--T-----A---------------A-A--A-
SIVgagDX..TCC-----T--C-----G---------------C-T--G-
 #721       .....................................
           WSYTCAGTWGAYGAACARATCCAGTGGATGTACMGWCARC
```

FIG. 4B

```
SIV gag    -------C--A--------------T---A-GA-------
SIVgagDX..-------G--C--------------C---C-TC-------
 #761       .....................................
           AGAACCCSATMCCAGTAGGCAACATYTACMGKMGATGGAT

SIV gag    ---A-----GT----A--A--T--CA-A-----T-----A
SIVgagDX..---G-----TC----G--G--C--TC-T-----C-----G
 #801       .....................................
           CCARCTGGGKYTGCARAARTGYGTYMGWATGTAYAACCCR

SIV gag    --A-----------------------------------
SIVgagDX..--C-----------------
 #841       .....................................
           ACMAACATTCTAGATGTAAAACAAGGGCCAAAAGAGCCAT

SIV gag    -----------------------------------------
 #881       .....................................
           TTCAGAGCTATGTAGACAGGTTCTACAAAAGTTTAAGAGC

SIV gag    -----------------------------------------
 #921       .....................................
           AGAACAGACAGATGCAGCAGTAAAGAATTGGATGACTCAA

SIV gag    -----------------------------------------
 #961       .....................................
           ACACTGCTGATTCAAAATGCTAACCCAGATTGCAAGCTAG

SIV gag    -----------------------------------------
 #1001      .....................................
           TGCTGAAGGGGCTGGGTGTGAATCCCACCCTAGAAGAAAT

SIV gag    -----------------------------------------
 #1041      .....................................
           GCTGACGGCTTGTCAAGGAGTAGGGGGGCCGGGACAGAAG

SIV gag    -----------------------------------------
 #1081      .....................................
           GCTAGATTAATGGCAGAAGCCCTGAAAGAGGCCCTCGCAC

SIV gag    -----------------------------------------
 #1121      .....................................
           CAGTGCCAATCCCTTTTGCAGCAGCCCAACAGAGGGGACC

SIV gag    -----------------------------------------
 #1161      .....................................
           AAGAAAGCCAATTAAGTGTTGGAATTGTGGGAAAGAGGGA
```

FIG. 4C

```
SIV gag  ----------------------------------------
#1201    ........................................
         CACTCTGCAAGGCAATGCAGAGCCCCAAGAAGACAGGGAT

SIV gag  ----------------------------------------
#1241    ........................................
         GCTGGAAATGTGGAAAAATGGACCATGTTATGGCCAAATG

SIV gag  ----------------------------------------
#1281    ........................................
         CCCAGACAGACAGGCGGGTTTTTTAGGCCTTGGTCCATGG

SIV gag  ----------------------------------------
#1321    ........................................
         GGAAAGAAGCCCCGCAATTTCCCCATGGCTCAAGTGCATC

SIV gag  ----------------------------------------
#1361    ........................................
         AGGGGCTGATGCCAACTGCTCCCCCAGAGGACCCAGCTGT

SIV gag  ----------------------------------------
#1401    ........................................
         GGATCTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG

SIV gag  ----------------------------------------
#1441    ........................................
         AGAGAAAAGCAGAGAGAAAGCAGAGAGAAGCCTTACAAGG

SIV gag  ----------------------------------------
#1481    ........................................
         AGGTGACAGAGGATTTGCTGCACCTCAATTCTCTCTTTGG

SIV gag  ----------------------------------------
#1521    ........................................
         AGGAGACCAGTAG
```

FIG. 4D

42/45

```
                                                BsrGI (37)
   1 CCTGGCCATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCATGTTGA
  81 CATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTAC
 161 ATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCA
 241 TAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
 321 GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGAC
                                                SnaBI (432)
 401 CTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACA
 481 TCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGG
 561 CACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTG
 641 GGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATA
                                                SacII (746)
 721 GAAGACACCGGGACCGATCCAGCCTCCGCGGGCCGCGCTAAGTATGGGATGTCTTGGGAATCAGCTGCTTATCGCCATCT

                                                1▶MetGlyCysLeuGlyAsnGlnLeuLeuIleAlaIleL
 801 TGCTTTTAAGTGTCTATGGGATCTATTGTACTCTATATGTCACAGTCTTTTATGGTGTACCAGCTTGGAGGAATGCGACA

                                                13▶euLeuLeuSerValTyrGlyIleTyrCysThrLeuTyrValThrValPheTyrGlyValProAlaTrpArgAsnAlaThr
 881 ATTCCCCTCTTTTGTGCAACCAAGAATAGGGATACTTGGGGAACAACTCAGTGCCTACCAGATAATGGTGATTATTCAGA

                                                40▶IleProLeuPheCysAlaThrLysAsnArgAspThrTrpGlyThrThrGlnCysLeuProAspAsnGlyAspTyrSerGl
 961 AGTGGCCCTTAATGTTACAGAAAGCTTTGATGCCTGGAATAATACAGTCACAGAACAGGCAATAGAGGATGTATGGCAAC

                                                66▶uValAlaLeuAsnValThrGluSerPheAspAlaTrpAsnAsnThrValThrGluGlnAlaIleGluAspValTrpGlnL
1041 TCTTTGAGACCTCAATAAAGCCTTGTGTAAAATTATCCCCATTATGCATTACTATGAGATGCAATAAAAGTGAGACAGAT

                                                93▶euPheGluThrSerIleLysProCysValLysLeuSerProLeuCysIleThrMetArgCysAsnLysSerGluThrAsp
1121 AGATGGGGATTGACAAAATCAATAACAACAACAGCATCAACAACATCAACGACAGCATCAGCAAAAGTAGACATGGTCAA

                                                120▶ArgTrpGlyLeuThrLysSerIleThrThrThrAlaSerThrThrSerThrThrAlaSerAlaLysValAspMetValAs
1201 TGAGACTAGTTCTTGTATAGCCCAGGATAATTGCACAGGCTTGGAACAAGAGCAAATGATAAGCTGTAAATTCAACATGA

                                                146▶nGluThrSerSerCysIleAlaGlnAspAsnCysThrGlyLeuGluGlnGluGlnMetIleSerCysLysPheAsnMetT
                                                                                PstI (1329)
1281 CAGGGTTAAAAAGAGACAAGAAAAAAGAGTACAATGAAACTTGGTACTCTGCAGATTTGGTATGTGAACAAGGGAATAAC

                                                173▶hrGlyLeuLysArgAspLysLysLysGluTyrAsnGluThrTrpTyrSerAlaAspLeuValCysGluGlnGlyAsnAsn
1361 ACTGGTAATGAAAGTAGATGTTACATGAACCACTGTAACACTTCTGTTATCCAAGAGTCTTGTGACAAACATTATTGGGA

                                                200▶ThrGlyAsnGluSerArgCysTyrMetAsnHisCysAsnThrSerValIleGlnGluSerCysAspLysHisTyrTrpAs
1441 TGCTATTAGATTTAGGTATTGTGCACCTCCAGGTTATGCTTTGCTTAGATGTAATGACACAAATTATTCAGGCTTTATGC

                                                226▶pAlaIleArgPheArgTyrCysAlaProProGlyTyrAlaLeuLeuArgCysAsnAspThrAsnTyrSerGlyPheMetP
1521 CTAAATGTTCTAAGGTGGTGGTCTCTTCATGCACAAGGATGATGGAGACACAGACTTCTACTTGGTTTGGCTTTAATGGA

                                                253▶roLysCysSerLysValValValSerSerCysThrArgMetMetGluThrGlnThrSerThrTrpPheGlyPheAsnGly
1601 ACTAGAGCAGAAAATAGAACTTATATTTACTGGCATGGTAGGGATAATAGGACTATAATTAGTTTAAATAAGTATTATAA

                                                280▶ThrArgAlaGluAsnArgThrTyrIleTyrTrpHisGlyArgAspAsnArgThrIleIleSerLeuAsnLysTyrTyrAs
1681 TCTAACAATGAAATGTAGAAGACCAGGAAATAAGACAGTTTTACCAGTCACCATTATGTCTGGATTGGTTTTCCACTCAC

                                                306▶nLeuThrMetLysCysArgArgProGlyAsnLysThrValLeuProValThrIleMetSerGlyLeuValPheHisSerG
                                                                XcmI (1778)
1761 AACCAATCAATGATAGGCCAAAGCAGGCATGGTGTTGGTTTGGAGGAAAATGGAAGGATGCAATAAAAGAGGTGAAGCAG

                                                333▶InProIleAsnAspArgProLysGlnAlaTrpCysTrpPheGlyGlyLysTrpLysAspAlaIleLysGluValLysGln
1841 ACCATTGTCAAACATCCCAGGTATACTGGAACTAACAATACTGATAAAATCAATTTGACGGCTCCTGGAGGAGGAGATCC

                                                360▶ThrIleValLysHisProArgTyrThrGlyThrAsnAsnThrAspLysIleAsnLeuThrAlaProGlyGlyGlyAspPr
1921 GGAAGTTACCTTCATGTGGACAAATTGCAGAGGAGAGTTCCTCTACTGTAAAATGAATTGGTTTCTAAATTGGGTAGAAG

                                                386▶oGluValThrPheMetTrpThrAsnCysArgGlyGluPheLeuTyrCysLysMetAsnTrpPheLeuAsnTrpValGluA
2001 ATAGGAATACAGCTAACCAGAAGCCAAAGGAACAGCATAAAAGGAATTACGTGCCATGTCATATTAGACAAATAATCAAC

                                                413▶spArgAsnThrAlaAsnGlnLysProLysGluGlnHisLysArgAsnTyrValProCysHisIleArgGlnIleIleAsn
```

FIG. 17A

43/45


PmlI (2134)
2081 ACTTGGCATAAAGTAGGCAAAAATGTTTATTTGCCTCCAAGAGAGGGAGACCTCACGTGTAACTCCACAGTGACCAGTCT
440 Thr Trp His Lys Val Gly Lys Asn Val Tyr Leu Pro Pro Arg Glu Gly Asp Leu Thr Cys Asn Ser Thr Val Thr Ser Le
2161 CATAGCAAACATAGATTGGATTGATGGAAACCAAACTAATATCACCATGAGTGCAGAGGTGGCAGAACTGTATCGATTGG
466 u Ile Ala Asn Ile Asp Trp Ile Asp Gly Asn Gln Thr Asn Ile Thr Met Ser Ala Glu Val Ala Glu Leu Tyr Arg Leu G
2241 AATTGGGAGATTATAAATTAGTAGAGATCACTCCAATTGGCTTGGCCCCCACAGATGTGAAGAGGTACACTACTGGTGGC
493 l Leu Gly Asp Tyr Lys Leu Val Glu Ile Thr Pro Ile Glu Leu Ala Pro Thr Asp Val Lys Arg Tyr Thr Thr Gly Gly
BspMI (2378)
2321 ACCTCAAGAAATAAAAGAGGGGTCTTTGTGCTAGGGTTCTTGGGTTTTCTCGCAACGGCAGGTTCTGCAATGGGAGCCGC
520 Thr Ser Arg Asn Lys Arg Gly Val Phe Val Leu Gly Phe Leu Gly Phe Leu Ala Thr Ala Gly Ser Ala Met Gly Ala Ala
2401 CAGCCTGACCCTCACGGCACAGTCCCGAACTTTATTGGCTGGGATAGTCCAACAGCAGCAACAGCTGTTGGACGTGGTCA
546 a Ser Leu Thr Leu Thr Ala Gln Ser Arg Thr Leu Leu Ala Gly Ile Val Gln Gln Gln Gln Leu Leu Asp Val Val L
Eam1105I (2502)
2481 AGAGACAACAAGAATTGTTGCGACTGACCGTCTGGGGAACAAAGAACCTCCAGACTAGGGTCACTGCCATCGAGAAGTAC
573 ys Arg Gln Gln Glu Leu Leu Arg Leu Thr Val Trp Gly Thr Lys Asn Leu Gln Thr Arg Val Thr Ala Ile Glu Lys Tyr
2561 TTAAAGGACCAGGCGCAGCTGAATGCTTGGGGATGTGCGTTTAGACAAGTCTGCCACACTACTGTACCATGGCCAAATGC
600 Leu Lys Asp Glu Ala Gln Leu Asn Ala Trp Gly Cys Ala Phe Arg Gln Val Cys His Thr Thr Val Pro Trp Pro Asn Al
2641 AAGTCTAACACCAAAGTGGAACAATGAGACTTGGCAAGAGTGGGAGCGAAAGGTTGACTTCTTGGAAGAAAATATAACAG
626 a Ser Leu Thr Pro Lys Trp Asn Asn Glu Thr Trp Gln Glu Trp Glu Arg Lys Val Asp Phe Leu Glu Glu Asn Ile Thr A
2721 CCCTCCTAGAGGAGGCACAAATTCAACAAGAGAAGAACATGTATGAATTACAAAAGTTGAATAGCTGGGATGTGTTTGGC
653 l Ala Leu Leu Glu Glu Ala Gln Ile Glu Gln Gln Glu Lys Asn Met Tyr Glu Leu Gln Lys Leu Asn Ser Trp Asp Val Phe Gly
2801 AATTGGTTTGACCTTGCTTCTTGGATAAAGTATATACAATATGGAGTTTATATAGTTGTAGGAGTAATACTGTTAAGAAT
680 Asn Trp Phe Asp Leu Ala Ser Trp Ile Lys Tyr Ile Glu Gln Tyr Gly Val Tyr Ile Val Val Gly Val Ile Leu Leu Arg Il
2881 AGTGATCTATATAGTACAAATGCTAGCTAAGTTAAGGCAGGGGTATAGGCCAGTGTTCTCTTCCCCACCCTCTTATTTCC
706 e Val Ile Leu Tyr Ile Val Gln Met Leu Ala Lys Leu Arg Gln Gly Tyr Arg Pro Val Phe Ser Ser Pro Pro Ser Tyr Phe G
PpuMI (2979)
2961 AGCAGACCCATATCCAACAGGACCCGGCACTGCCAACCAGAGAAGGCAAAGAAAGAGACGGTGGAGAAGGCGGTGGCAAC
733 l Gln Gln Thr His Ile Glu Gln Gln Asp Pro Ala Leu Pro Thr Arg Glu Gly Lys Glu Arg Asp Gly Gly Glu Gly Gly Asn
3041 AGCTCCTGGCCTTGGCAGATAGAATATATCCACTTTCTTATTCGTCAGCTTATTAGACTCTTGACTTGGCTATTCAGTAA
760 Ser Ser Trp Pro Trp Gln Ile Glu Tyr Ile His Phe Leu Ile Arg Gln Leu Ile Arg Leu Leu Thr Trp Leu Phe Ser As
3121 CTGTAGGACTTTGCTATCGAGAGTATACCAGATCCTCCAACCAATACTCCAGAGGCTCTCTGCGACCCTACAGAGGATTC
786 n Cys Arg Thr Leu Leu Ser Arg Val Tyr Gln Ile Ile Leu Gln Pro Ile Leu Gln Arg Leu Ser Ala Thr Leu Gln Arg Ile A
Bsu36I (3208)
3201 GAGAAGTCCTCAGGACTGAACTGACCTACCTACAATATGGGTGGAGCTATTTCCATGAGGCGGTCCAGGCCGTCTGGAGA
813 r g Glu Val Leu Arg Thr Glu Leu Thr Tyr Leu Gln Tyr Gly Trp Ser Tyr Phe His Glu Ala Val Gln Ala Val Trp Arg
3281 TCTGCGACAGAGACTCTTGCGGGCGCGTGGGGAGACTTATGGGAGACTCTTAGGAGAGGTGGAAGATGGATACTCGCAAT
840 Ser Ala Thr Glu Thr Leu Ala Gly Ala Trp Gly Asp Leu Trp Glu Thr Leu Arg Arg Gly Gly Ala Arg Trp Ile Leu Ala Il
BamHI (3418)
EcoRI (3412)
3361 CCCCAGGAGGATTAGACAAGGGCTTGAGCTCACTCTCTTGTGAGGGACAGAGAATTCGGATCCactagttctagaCTCGA
866 e Pro Arg Arg Ile Arg Gln Gly Leu Glu Leu Thr Leu Leu• • •
Eco47III (3457)
3441 GGGGGGGCCCGGTACGAGCGCTTAGCTAGCTAGAGACCACCTCCCCTGCGAGCTAAGCTGGACAGCCAATGACGGGTAAG
3521 AGAGTGACATTTTTCACTAACCTAAGACAGGAGGGCCGTCAGAGCTACTGCCTAATCCAAAGACGGGTAAAAGTGATAAA
BstEII (3673)
3601 AATGTATCACTCCAACCTAAGACAGGCGCAGCTTCCGAGGGATTTGTCGTCTGTTTTATATATATTTAAAAGGGTGACCT

FIG. 17B


BsaBI (3740)
SphI (3948)
KpnI (3976)
BstXI (4060)
XmnI (4293)
StuI (5368)
BsrFI (6036)
SspI (6067)
SmaI (6118)
NruI (6335)
PvuI (5993)
SgfI (5992)
DraIII (6523)

FIG. 17C

```
6641  TTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAG
6721  GCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTC
6801  TGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTAT
6881  GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCG
6961  CATCAGATTGGCTATTGG
```

FIG. 17D